LOCAL DUALITY AND STRUCTURES OF NONLINEAR PROGRAMS IN HILBERT SPACES

BY

YUEN CHUNG MAN

A

Thesis

submitted to

The Graduate School of

The Chinese University of Hong Kong

(Division of Mathematics)

In Partially Fulfillment

of the Requirements for the Degreee of

Master of Philosophy (M.Phil.)

May, 1985

Acknowledgement

I am greatly indebted to my supervisor, Dr. W.L. Chan, for his valuable advice and encouragement during my graduate study and in preparation of this thesis. I wish to thank Dr. L.F. Ho and my colleagues for their constructive suggestions.

Abstract

The aim of this thesis is to study local duality of nonlinear programmes in a Hilbert space H. In chapter one we collect some useful results and definitions. The first part of chapter two is concerned with the conjugate decomposition of a element of H with respect to a closed convex cone K. Next, basing on this result, we give an extension of the finite dimensional results of Han and Mangasarian characterizing positive definite matrices to positive definite operators on a Hilbert space. In chapter three we formulate geometrically meaningful second order necessary and sufficient optimality conditions for the dual problem of the Wolfe type and study the relationship to the corresponding conditions for the primal. Then we generalize the work of Fujiwara, Han and Mangasarian, deriving a characterization of simultaneous local solutions of the primal and dual problem. We present, in chapter four a modification of the work of Maurer and Zowe, developing first and second order necessary and sufficient conditions for optimality for nonlinear programmes in infinite dimensional spaces.

Contents

Chapter 1. Preliminary Results

In this chapter, we record some important notations, definitions and theorems which we shall use later on. (c.f. [1], pp. 16-30).

Basic Notation

X^*	the dual space of space X.
$L(X,Y)$	the space of continuous linear mappings of the space X into space Y.
$L(X)$	the space of continuous linear mappings of the space X into itself.
$\ker \Lambda$	the kernel of linear operator $\Lambda \in L(X,Y)$.
$\operatorname{Im} \Lambda$	the image of linear operator $\Lambda \in L(X,Y)$.
Λ^*	the adjoint operator of linear operator $\Lambda \in L(X,Y)$.
$F : X \to Y$	the mapping F from X into Y.
x^*h or $x^*(h)$	the mapping x^* acts on h.
$S^\circ = \{x^* \in X^* \mid x^*(s) \leq 0 \text{ for all } s \text{ in } S \subseteq X\}$	the polar cone of S.
$S^+ = \{x^* \in X^* \mid x^*(s) \geq 0 \text{ for all } s \text{ in } S \subseteq X\}$	the positive cone of S.
$L^\perp = \{x^* \in X^* \mid x^*(l) = 0 \text{ for all } l \text{ in } L \subseteq X\}$	the annihilator of the subspace L in X.
$\operatorname{int} S$	the interior of S.
$F'(x)$, $F''(x)$	the first and second Frechet derivative of the mapping F at x.
$F_{x_1}(x_1,x_2)$ $(F_{x_2}(x_1,x_2))$	the partial derivative with respect to x_1 (x_2) of the mapping $x_1 \to F(x_1,x_2)$ $(x_2 \to F(x_1,x_2))$.

C^1 the class of all continuous Frechet differentiable functions on some open set in X.

$\|x\|$ the norm of element x in a Banach space.

Definition 1.1.

Let X and Y be Banach Spaces. Let F be a mapping from X into Y. Let $x_o \in X$ and F be Frechet differentiable at x_o. F is called regular at x_o if

$$F'(x_o)X = Y.$$

Theorem 1.2. (Implicit Function Theorem).

Let X, Y and Z be Banach spaces. Let U be a neighborhood of a point (x_o,y_o) in the Cartesian product X x Y, and let $F : U \to Z$ be a mapping of class C^1. Further, assume that $F(x_o,y_o) = 0$, and that the partial derivative $F_y(x_o,y_o) : Y \to Z$ is a linear homeomorphism.

Then $\exists\ \varepsilon > 0$ and $\delta > 0$, and a mapping $x \to y(x)$ of the ball $U(x_o,\delta) \subset X$ into the ball $U(y_o,\varepsilon) \subset Y$ such that

(a) the relations $F(x,y) = 0$ and $y = y(x)$ are equivalent on the set $U(x_o,\delta) \times U(y_o,\varepsilon)$;

(b) $y(x)$ is of class C^1 and for any $x \in U(x_o,\delta)$

$$y'(x) = -[\, F_y(x,y(x))]^{-1} \circ\ F_x(x,y(x)).$$

Definition 1.3.

Let M be a subset of a Banach space X. Let $x_o \in M$.

$$T(M,x_o) := \Big\{ d \in X \mid \exists\ \varepsilon > 0 \text{ and a mapping } t \to r(t) : [0,\varepsilon] \to X, \text{ s.t. } x_o + td + r(t) \in M\ \forall t \in [0,\varepsilon] \text{ and } \|r(t)\|/t \to 0 \text{ as } t \to 0 \Big\}.$$

$$N(x_o) := T(M,x_o)^o.$$

$T(M,x_o)$ is called the <u>tangent cone</u> of M at x_o.

$N(x_o)$ is called the <u>polar cone</u> (or normal cone) of $T(M,x_o)$.

If $T(M,x_o)$ is a subspace of X then $T(M,x_o)$ is called the <u>tangent space</u> of M at x_o and $N(x_o)$ becomes <u>annihilator</u> of $T(M,x_o)$, i.e. $N(x_o) = T(M,x_o)^{\perp}$.

Theorem 1.4. (Ljusternik's Theorem).

Let X and Y be Banach spaces. Let U be a neighborhood of a point $x_o \in X$, and let F be a Frechet differentiable mapping of U into Y. Assume that F is regular at x_o and that its derivative is continuous at this point (in uniform operator topology of the space $L(X,Y)$). Then the tangent space to the set

$$M = \{ x \in U \mid F(x) = F(x_o) \}$$

at the point x_o is

$$T(M,x_o) = \ker F'(x_o).$$

Moreover, if the assumptions of the theorem are satisfied then there exists a neighborhood $U' \subset U$ of the point x_o, a number $K > 0$, and a mapping $\xi \to x(\xi)$ of the set U' into X s.t. for all $\xi \in U'$

$$F(\xi + x(\xi)) = F(x_o)$$

and

$$\| x(\xi) \| \leq K \| F(\xi) - F(x_o) \| .$$

Lemma 1.5. (lemma on the annihilator).

Let X and Y be Banach spaces, and let $\Lambda : X \to Y$ be a continuous linear operator s.t. $\operatorname{Im} \Lambda = Y$. Then

$$(\ker \Lambda)^{\perp} = \operatorname{Im} \Lambda^{*},$$

i.e. the annihilator of the kernel is equal to the range of the adjoint operator.

1.6 Remark :

Assume the conditions of the theorem 1.4 hold. Then

$$T(M,x_o) = \ker F'(x_o).$$

By lemma 1.5, the annihilator $N(x_o)$ of $T(M,x_o)$ is

$$\begin{aligned} N(x_o) &= T(M,x_o)^{\perp} \\ &= \ker F'(x_o)^{\perp} \\ &= \left\{ x^* \in X^* \mid \exists\, y^* \in Y^*, \text{ s.t. } x^* = F'(x_o)^* y^* \right\}, \end{aligned}$$

i.e. $N(x_o) = F'(x_o)^* Y^*$.

Chapter 2. Characterization of Positive Definite Operators on a Hilbert Space

In [5], the authors proved that an n x n matrix A is positive definite iff it is positive definite on some closed convex cone K in $\mathbb{R}^n$ and $(A + A^T)^{-1}$ exists and is positive semidinfinite on the polar cone K^o. In this chapter, we generalize these results to a Hilbert space operator T by considering the conjugate decomposition of the element in the Hilbert space.

Throughout this chapter, $(H, \langle \cdot , \cdot \rangle)$ denotes a real Hilbert space, T is a bounded linear operator on H and K is a closed convex cone in H.

2.1 Definitions :

(a) T is called positive semidefinite (PSD) on K if $\langle Tx,x \rangle \geq 0 \ \forall x \in K$

(b) T is called positive (PS) on K if $\langle Tx,x \rangle > 0 \ \forall x \in K,\ x \neq 0$.

(c) T is called positive definite on K if $\exists\, \alpha > 0$ s.t. $\langle Tx,x \rangle \geq \alpha \|x\|^2$ $\forall\, x \in K$. (T may be called coercive on K and α is called a coercive constant).

(d) $K^T := \{ y \in H \mid \langle (T + T^*)y,x \rangle \leq 0,\ \forall\, x \in K \}$, where T^* is the adjoint operator of T. K^T is called the conjugate cone in H with respect to K and T.

(e) Let $h \in H$. h is said to have a <u>conjugate decomposition</u> with respect to K and T if $\exists\, x \in K,\ y \in K^T$ s.t. $h = x + y$, and $\langle (T + T^*)y,x \rangle = 0$.

Theorem 2.2.

Let $T \in L(H)$ and K be any closed convex cone in the Hilbert space $(H,\langle\cdot,\cdot\rangle)$.

Assume that the function $g(x) = \langle Tx,x\rangle$ is weakly lower semicontinuous (WLSC) on K and that T is positive definite on K. Then

(i) any vector h in H has a conjugate decomposition with respect to K and T;

(ii) If T is positive on the linear hull of K then the decomposition of (i) is unique.

Proof :

(i) Since T is positive definite on K, i.e. $\exists\, \alpha > 0$ s.t. $\langle Tx,x\rangle \geq \alpha \|x\|^2 \ \forall x \in K$.

Let $h \in H$, and let

$$S := \left\{ x \in K \mid \|x\| \leq \frac{\|(T + T^*)h\|}{\alpha} \right\}.$$

Since 0 belongs to K, $S \neq \emptyset$.

Let $f(x) = \langle T(x-h),x-h\rangle$ and consider the following problems :

$$(Q.1) \qquad \min\left\{ f(x) \mid x \in K \right\},$$

$$(Q.2) \qquad \min\left\{ f(x) \mid x \in S \right\}.$$

Any solution of (Q.2) solves (Q.1) since for any $x \in K\backslash S$

$$\begin{aligned} f(x) &= \langle(T(x-h),x-h\rangle \\ &= \langle Tx,x\rangle - \langle Tx,h\rangle - \langle Th,x\rangle + \langle Th,h\rangle \\ &= \langle Tx,x\rangle - \langle(T + T^*)h,x\rangle + \langle Th,h\rangle \\ &\geq \alpha\|x\|^2 - \|(T + T^*)h\|\,\|x\| + f(0) \\ &= (\alpha\|x\| - \|(T + T^*)h\|)\,\|x\| + f(0) \end{aligned}$$

$$> f(0)$$
$$\geq \min\left\{f(x) \mid x \in S\right\}.$$

Since S, being norm bounded ,closed and convex, is weakly compact and f(x) is weakly lower semicontinuous on K, as $\langle Tx,x\rangle$ is WLSC on K, $\exists\ \bar{x} \in K$ s.t.

$$f(\bar{x}) = \min_{x\in S} f(x) = \min_{x\in K} f(x) .$$

Since $\bar{x}$ minimize f over K, $f'(\bar{x})(y - \bar{x}) \geq 0\ \forall y \in K$,

i.e. $f'(\bar{x})(y - \bar{x}) = \langle (T + T^*)(\bar{x} - h), y - \bar{x}\rangle \geq 0\ \forall y \in K$.

Since K is a cone and $\bar{x} \in K$, 0 and $2\bar{x} \in K$. Therefore,

$$\langle (T + T^*)(\bar{x} - h),\bar{x}\rangle \leq 0$$

and

$$\langle (T + T^*)(\bar{x} - h),\bar{x}\rangle \geq 0,$$

i.e. $\langle (T + T^*)(\bar{x} - h),\bar{x}\rangle = 0$ and $\langle (T + T^*)(\bar{x} - h),y\rangle \geq 0\ \forall\ y\in K$.

Let $-\bar{y} = \bar{x} - h$, then

$$h = \bar{x} + \bar{y},\ \bar{x} \in K,\ \bar{y} \in K^T \text{ and } \langle (T + T^*)\bar{y},\bar{x}\rangle = 0.$$

Hence h has a conjugate decomposition.

(ii) Now we assume that T is positive on the linear hull of K and go to prove the uniqueness of the conjugate decomposition obtained in part (i).

Suppose $h \in H$ and $h = x + y = \bar{x} + \bar{y}$, where $x,\ \bar{x} \in K$, $y,\ \bar{y} \in K^T$ s.t.

$$\langle (T + T^*)y,x\rangle = \langle (T + T^*)\bar{y},\bar{x}\rangle = 0.$$

Since $\bar{x} - x = y - \bar{y}$,

$$\begin{aligned} &\langle T(\bar{x} - x),\bar{x} - x\rangle \\ &= \langle (T(\bar{x} - x),y - \bar{y}\rangle \\ &= \langle (T(y - \bar{y}),\bar{x} - x\rangle \\ &= \langle (T^*(y - \bar{y}),\bar{x} - x\rangle \\ &= \tfrac{1}{2}\langle (T + T^*)(y- \bar{y}),\bar{x} - x\rangle \end{aligned}$$

$$= \tfrac{1}{2}\left\{ \langle (T + T^*)y,\bar{x}\rangle - \langle (T + T^*)\bar{y},\bar{x}\rangle - \langle (T + T^*)y,x\rangle + \langle (T + T^*)\bar{y},x\rangle \right\}$$
$$= \tfrac{1}{2}\left\{ \langle (T + T^*)y,\bar{x}\rangle + \langle (T + T^*)\bar{y},x\rangle \right\}$$
$$\leq 0.$$

Since T is positive on the linear hull of K, the above inequality holds only when $x = \bar{x}$. Hence the conjugate decomposition is unique.

Lemma 2.3. (c.f. preposition 3.2, in [7]).

If T is PSD on H, then $Q(x) = \langle Tx,x\rangle$ is WLSC on H.

Now, we consider a particular case. Let T be the identity operator on H. Then the conjugate cone K^T becomes the polar cone K^o of K. I is positive on the linear hull of K. Therefore, we obtain the following corollary.

Corollary 2.4.

Let K be a closed convex cone in H. Then for any element h in H, $\exists$ unique $x_h \in K$ and $y_h \in K^o$ s.t.

(i) $h = x_h + y_h$ and $\langle x_h, y_h\rangle = 0$,

(ii) $\min_{\hat{x} \in K} \|\hat{x} - h\| = \|x_h - h\|$,

(iii) $\|h\|^2 = \|x_h\|^2 + \|y_h\|^2$, and

(iv) If $h = x_1 + y_1$, $x_1 \in K$, $y_1 \in K^o$ then $\|x_1\| \geq \|x_h\|$, $\|y_1\| \geq \|y_h\|$.

Proof :

By lemma 2.3, $\langle Ix,x\rangle$ is WLSC on K. Then (i), (ii), (iii) are obtained from the proof of theorem 2.2. It remains to prove part (iv). Since

$$\|x_h - h\| = \min_{\tilde{x} \in K} \|\tilde{x} - h\|,$$

$$\|y_h\| \leq \min_{\tilde{x} \in K} \|\tilde{x} - h\|$$

$$\leq \|x_1 - h\|$$

$$= \|y_1\|.$$

As K is closed, $K^{oo} = K$.(c.f. p.383, [13]). By the symmetric position of K and K^o, we have

$$\|x_h\| \leq \|x_1\|.$$

Remark : this decomposition is called the <u>polar decomposition</u>.

Since the polar decomposition is unique, we can define a mapping $P^K : H \to K$. We characterize some properties about this mapping as follow :

Definition 2.5.

Define $P^K : H \to K$ as $P^K(h) = x_h$.

P^K is called the projection of H on the closed convex cone K and x_h is called the image of h under P^K.

Corollary 2.6.

$\forall\, h' \in K$, $h' = P^K(h)$ iff $h - h' \in K^o$ and $\langle h', h-h' \rangle = 0$.

Proof :

The necessity directly follows corollary 2.4(i).

Let $y \in K$. Then we consider the following :

$$\| y - h \|^2 = \| y - h' + h' - h \|^2$$

$$= \langle y - h' + h' - h \ , \ y - h' + h' - h\rangle$$
$$= \| y - h'\|^2 + \| h' - h \|^2 + 2\langle y,h'-h\rangle - 2\langle h',h'-h\rangle$$
$$\geq \| h' - h\|^2 \ .$$

By the corollary 2.4 (ii), $h' = x_h = P^K(h)$.

Corollary 2.7.

For any $h \in H$, the followings hold :

(i) $P^K(-h) = -P^{-K}(h)$.

(ii) $P^{-K^\circ}(h) = h + P^K(-h)$.

(iii) $P^K(\lambda h) = \lambda P^K(h) \ \ \forall \lambda \geq 0$.

(iv) $\forall \bar{h} \in K$, $\bar{h} = P^K(h)$ holds iff $\| \bar{h} \| = \min\limits_{d \in -K^\circ + h} \| d \|$.

Proof :

(i) By corollary 2.4(i), we have

$$h = P^{-K}(h) + P^{-K^\circ}(h)$$
$$\Rightarrow h - P^{-K}(h) \in -K^\circ \text{ and } \langle h - P^{-K}(h), P^{-K}(h)\rangle = 0$$
$$\Rightarrow P^{-K}(h) - h \in K^\circ \text{ and } \langle h - P^{-K}(h), P^{-K}(h)\rangle = 0$$
$$\Rightarrow (-h) - (-P^{-K}(h)) \in K^\circ \text{ and } \langle -h - (-P^{-K}(h)), -P^{-K}(h)\rangle = 0.$$

Hence, by corollary 2.6, we have $P^K(-h) = -P^{-K}(h)$.

(ii) By corollary 2.4(i), we have

$$h = P^{-K}(h) + P^{-K^\circ}(h)$$
$$= -P^K(-h) + P^{-K^\circ}(h),$$

i.e. $\quad P^{-K^\circ}(h) = h + P^K(-h)$.

(iii) Since $\lambda h - \lambda P^K(h) = \lambda(h - P^K(h)) \in K^\circ \ \forall \lambda \geq 0$ and

$$\langle \lambda h - \lambda P^K(h), \lambda P^K(h)\rangle$$
$$= \lambda^2 \langle h - P^K(h), P^K(h)\rangle$$
$$= 0 \ ,$$

by the corollary 2.6, we have

$$P^K(\lambda h) = \lambda P^K(h).$$

(iv) By corollary 2.4(i) and (ii), we have

$$\| P^K(h) \| = \| h - P^{K^\circ}(h) \|$$

$$= \min_{y \in K^\circ} \| h - y \|$$

$$= \min_{d \in -K^\circ + h} \| d \|.$$

Hence, $P^K(h) = \bar{h}$ iff $\| \bar{h} \| = \min_{d \in -K^\circ + h} \| d \|.$

Here, we also note the following properties of $P^K(h)$. (c.f. lemma 2.5, 2.7 and corollary 2.6 in [8]).

2.8 $\| P^K(h) \| \leq \| h \|$, $\forall\, h \in H$.

2.9 $\| P^K(h_1) - P^K(h_2) \| \leq \| h_1 - h_2 \|$, $h_1, h_2 \in H$.

2.10 $\| P^K(h_1 + h_2 - d^*) \| \leq \| P^K(h_1) + P^K(h_2) \|$, $h_1, h_2 \in H$, and $d^* \in K^\circ$.

2.11 The functional $\| P^K(h) \|$ is convex.

Let $q(h) = \frac{1}{2} \| P^{-K^\circ}(h) \|^2$.

2.12 $q(\lambda h_1 + (1-\lambda) h_2 - d) \leq \lambda q(h_1) + (1-\lambda) q(h_2)$, $h_1, h_2 \in H$, $d \in K$ and $\lambda \in (0,1)$. In particular, $q(h)$ is convex.

2.13 The functional $q(h)$ is Frechet differentiable with the derivative

$$q'(h) = P^{K^\circ}(h).$$

When K is a closed subspace of H, K° becomes the orthogonal complement $K^\perp$ of K and the decomposition becomes the orthogonal decomposition.

Now, we state and prove the main theorem in this chapter.

Theorem 2.14.

Let $T \in L(H)$ and let K be a closed convex cone in H. Then the following conditions are equivalent :

(i) T is positive definite on H.

(ii) T is positive definite on K and K^T, and $\langle Tx,x\rangle$ is WLSC on K.

(iii) T is positive definite on K, $(T + T^*)$ is invertible and $(T + T^*)^{-1}$ is positive definite on K^0, and $\langle Tx,x\rangle$ is WLSC on K.

Proof :

(i) => (ii) :

The weakly lower semicontinuity of $\langle Tx,x\rangle$ is given by lemma 2.3. The other results trivially hold.

(ii) => (i) :

Since T is positive definite on K and $\langle Tx,x\rangle$ is WLSC on K, by theorem 2.2, $\forall h \in H$ has a conjugate decomposition

$$h = x + y,\ x \in K,\ y \in K^T,\ \langle (T + T^*)y,x\rangle = 0.$$

$$\begin{aligned}&\langle Th,h\rangle\\ &= \langle T(x + y),x + y\rangle\\ &= \langle Tx,x\rangle + \langle Ty,y\rangle + \langle (T + T^*)y,x\rangle\\ &= \langle Tx,x\rangle + \langle Ty,y\rangle.\end{aligned}$$

Since T is positive definite on K and K^T, T is positive definite on H.

(ii) => (iii) :

Since (i) <=> (ii), T is positive definite on H. Then $(T + T^*)$ is positive definite on H, i.e. $\langle (T + T^*)x,x\rangle \geq \alpha \|x\|^2$, for some $\alpha > 0$. Since $\frac{1}{\alpha}(T + T^*)$ is a self-adjoint bounded operator and satisfies $\frac{1}{\alpha}(T + T^*) \geq I$, by theorem 2-1 in [3, p.108], $\frac{1}{\alpha}(T + T^*)$ is invertible and so is $(T + T^*)$. It remains to prove that $(T + T^*)^{-1}$ is positive definite on K^o.

We claim that

$$(T + T^*)^{-1}K^o = K^T.$$

Let $y \in K^o$. Then $\forall\, x \in K$,

$$\begin{aligned} &\langle (T + T^*)(T + T^*)^{-1}y, x\rangle \\ &= \langle y, x\rangle \\ &\leq 0, \end{aligned}$$

i.e. $(T + T^*)^{-1}K^o \subseteq K^T$.

Let $y \in K^T$. Then $\forall\, x \in K$, $\langle (T + T^*)y, x\rangle \leq 0$. Therefore, $(T + T^*)y \in K^o$. It means that $y \in (T + T^*)^{-1}K^o$, i.e. $K^T \subseteq (T + T^*)^{-1}K^o$. Hence $(T + T^*)^{-1}K^o = K^T$.

Now, let $y \in K^o$.

$$\begin{aligned} &\langle (T + T^*)^{-1}y, y\rangle \\ &= \langle w, (T + T^*)w\rangle \quad \text{, where } (T + T^*)w = y \text{ and } w \in K^T \\ &= \langle (T + T^*)w, w\rangle \\ &= 2\langle Tw, w\rangle \\ &\geq 2\alpha \|w\|^2, \end{aligned}$$

since T is positive definite on K^T, i.e. $\langle Tw, w\rangle \geq \alpha \| w \|^2$, for some $\alpha > 0$ $\forall\, w \in K^T$. But

$$\begin{aligned} \| y \|^2 &= \| (T + T^*)w \|^2 \\ &\leq \| (T + T^*)\|^2 \|w\|^2. \end{aligned}$$

Thus, $$\langle (T + T^*)^{-1}y, y\rangle \geq \frac{2\alpha}{\| T + T^* \|^2} \| y \|^2, \ \forall\, y \in K^o.$$

Hence, $(T + T^*)$ is invertible and its inverse is positive definite on K^o.

(iii) => (i) :

Since T is positive definite on K and $\langle Tx,x\rangle$ is WLSC on K. By theorem 2.2, any vector h in H has a conjugate decomposition with respect to K and T, i.e.

$$h = x + y,\ x \in K,\ y \in K^T,\ \langle (T + T^*)y,x\rangle = 0.$$

Then

$$\begin{aligned}
&\langle Th,h\rangle \\
&= \langle Tx,x\rangle + \langle Ty,y\rangle + \langle (T + T^*)y,x\rangle \\
&= \langle Tx,x\rangle + \tfrac{1}{2}\langle (T + T^*)y,y\rangle \\
&= \langle Tx,x\rangle + \tfrac{1}{2}\langle (T + T^*)^{-1}w,w\rangle, \text{ where } (T + T^*)y = w \text{ and } w \in K^o, \\
&\geq \alpha \|x\|^2 + \tfrac{1}{2}\beta \| w \|^2, \text{ for some } \alpha, \beta > 0,
\end{aligned}$$

since T and $(T + T^*)^{-1}$ are positive definite on K and K^o respectively.

But

$$y = (T + T^*)^{-1}w,$$

i.e. $$\| y \|^2 \leq \|(T + T^*)^{-1}\|^2 \| w \|^2.$$

Therefore,

$$\begin{aligned}
&\langle Th,h\rangle \\
&\geq \alpha \| x \|^2 + \tfrac{1}{2}\frac{\beta}{\| (T + T^*)^{-1}\|^2}\| y \|^2 \\
&\geq \gamma(\| x \|^2 + \| y \|^2) \\
&\geq \tfrac{1}{2}\gamma \| h \|^2,
\end{aligned}$$

where $\gamma = \min\left\{\alpha , \tfrac{1}{2}\frac{\beta}{\|(T+ T^*)^{-1}\|^2}\right\} \geq 0.$

Hence the proof is completed.

Corollary 2.15.

Let $T \in L(H)$ and let K be a closed convex cone in H. Then the following conditions are equivalent :

(i) T is positive definite on H,

(ii) T is positive definite on K and K^T, and $\langle Tx,x\rangle$ is WLSC on K, and

(iii) T is positive definite on K, $(T + T^*)$ is invertible and $(T + T^*)^{-1}$ is positive on K^o, and $\langle Tx,x\rangle$ is WLSC on K.

Proof :

It suffices to prove that (iii) => (i).

From the proof of the part (iii) => (i) of the theorem 2.14, we know that T is positive on H. Then $T + T^*$ is positive on H. Since $T + T^*$ is invertible, $T + T^*$ is positive definite on H [14, theorem 2] and so is T.

We give a proposition used to check whether $(T + T^*)$ is invertible or not.

Proposition 2.16.

Let $T \in L(H)$ and let T be a self-adjoint operator.

If T is one to one, then T^{-1} exists on the range of T. Furthermore, if T^{-1} is bounded on the range of T, then T is invertible.

Proof :

The first statement is obviously true. Clearly T^{-1} is linear.

We claim that the range of T is dense in H. Let $f \in H$ and $\langle f,Tg\rangle = 0$ for all elements g in H. Then

$$0 = \langle f, Tg \rangle$$
$$= \langle Tf, g \rangle \ \forall g \in H,$$

$\Rightarrow Tf = 0$

$\Rightarrow f = 0$, since T is one to one,

$\Rightarrow$ the range of T is dense in H.

We have proved that T^{-1} exists on the range of T and the range of T is dense in H. It suffices to prove that the range of T is closed in H. Since T is bounded linear transformation and is defined on H, T is closed [4, theorem 16.2, p 263]. T^{-1} is also closed [4, theorem 16.3,p264]. Hence domain of T^{-1} is a closed subspace in H [4, theorem 16.4, p.264].

Remark : the inverse of the proposition is also true, which is given by the open mapping theorem.

Corollary 2.17.

Let $T \in L(H)$. $T + T^*$ is one to one and $(T + T^*)^{-1}$ is bounded on the range of $(T + T^*)$ iff $(T + T^*)$ is invertible.

Now, we consider the following example.

Example 2.18.

Let $l_2 = \{ (x_n) \mid x_i \in \mathbb{R} , \Sigma x_i^2 < \infty \}$.

Let E be a countable orthonormal basis in l_2. $E = \{ e_1, e_2, e_3, \dots \}$ where $\langle e_i, e_j \rangle = 0$ for $i \neq j$, $\| e_i \| = 1$, for all $i \in \mathbb{N}$.

Define $T : l_2 \to l_2$ as $Tx = \Sigma \alpha_i \langle x, e_i \rangle e_i$, where $(\alpha_i) \in l_2$, $\alpha_i > 0$ for all i and $\{\alpha_i\}$ decreases to zero.

Clearly, T is well-defined, since

$$\Sigma \alpha_i^2 \langle x, e_i\rangle^2 \leq \Sigma \alpha_i^2 \|x\|^2 \| e_i\|^2$$
$$= (\Sigma \alpha_i^2) \| x \|^2$$
$$< \infty,$$

i.e. $Tx \in l_2$.

We claim that T is one to one. Suppose $Tx = 0$. Then $\langle x,e_i\rangle = 0$ for all i , i.e. $x = 0$. Therefore T^{-1} exists in the range of T.

Since

$$\begin{aligned} &\langle Tx,y\rangle \\ &= \langle \Sigma \alpha_i \langle x,e_i\rangle e_i, \Sigma \langle y,e_j\rangle e_j\rangle \\ &= \Sigma \alpha_i \langle x,e_i\rangle \langle y,e_i\rangle \\ &= \langle x,Ty\rangle, \end{aligned}$$

i.e. T is 1-1 self-adjoint operator. Clearly T is PS on l_2 and thus $\langle Tx,x\rangle$ is WLSC on l_2.

Fixed $N > 0$. Let

$$K := \{(x_n) \in l_2 \mid x_i \geq 0,\ i = 1,2,\ldots,N,\ x_j = 0,\ j > N\}.$$

Clearly, K is a closed convex cone in l_2. Then

$$K^o = \{(x_n) \in l_2 \mid x_i \leq 0,\ i = 1,2,\ldots,N\} \neq \emptyset.$$

T is positive definite on K, i.e. $\langle Tx,x\rangle \geq \alpha \| x \|^2$, where

$$\alpha = \min \{\alpha_1, \alpha_2, \alpha_3, \cdots \alpha_N\} > 0.$$

T^{-1} is positive on $K^o \cap$ range of T. But T is not positive definite on l_2.

From this example, we find that the conditions in corollary 2.15, $(T + T^*)$ being invertible and its inverse being positive on K^o, cannot be reduced. But in finite dimensional case, these conditions can be reduced to $(T + T^*)$ being 1-1 and its inverse being positive semi-definite on K^o. (c.f. corollary 5 in [5]).

To end this chapter, we give the following corollaries :

Corollary 2.19.

Let $T \in L(H)$ and let T be self-adjoint. Let K be a closed subspace in H. Then the following conditions are equivalent :

(i) T is positive definite on H.

(ii) T is positive definite on K, and is positive definite on K^{T}, and $\langle Tx,x\rangle$ is WLSC on K .

(iii) T is positive on K, T is invertible and T^{-1} is positive definite (or positive) on $K^{\perp}$, and $\langle Tx,x\rangle$ is WLSC on K.

Corollary 2.20.

Let $T \in L(H)$ and let K be a closed convex cone in H.

If T is positive on K , $(T + T^{*})$ is invertible and $(T + T^{*})^{-1}$ is positive semidefinite on K^{o}, and $\langle Tx,x\rangle$ is WLSC on K, then T is positive semidefinite on H.

Corollary 2.21.

Let $T \in L(H)$ and let T be self-adjoint. Let K be a closed subspace in H.

If T is positive definite on K , T is invertible and T^{-1} is positive semidefinite on $K^{\perp}$, and $\langle Tx,x\rangle$ is WLSC on K then T is positive semidefinite on H.

Corollary 2.22.

Let $T \in L(H)$ and let K be a closed convex cone in H.

Assume that

(i) T is positive definite on K, and

(ii) T is positive (positive semidefinite) on K^{T}.

Then

T is positive (positive semidefinite) on H iff $\langle Tx,x\rangle$ is WLSC on H.

Remark : this corollary is also proved in [7,corollary 3.5].

Chapter 3. Local Duality of Nonlinear Programs

We consider the following problem :

(P) $\quad\quad \inf \quad J(x),$

subject to $G(x) \in K$,

where $J : X \to \mathbb{R}$,

$G : X \to Y$,

X, Y are Banach spaces,

K is a non-empty closed convex cone in Y.

We define its Wolfe dual problem as follow :

(D) $\quad\quad \sup J(x) - (l, G(x)),$

subject to $J'(x) - G'(x)^{*}l = 0$,

$l \in K^{+}$.

If x_o is a solution of (P), then under some regularity conditions, there exists a $l_o \in K^{+}$ such that $J'(x_o) - G'(x_o)^{*}l_o = 0$. Let

$$F(x) = J(x) - (l_o, G(x)).$$

We are going to show that the second order sufficient condition of problem (D), in the case K = 0, is equivalent to the invertibility of the second Frechet derivative of $F''(x_o)$ and the positive definiteness of $F''(x_o)$ on the annihilator $N(x_o)$ of the tangent space of (P) at x_o. Then we investigate the necessary and sufficient conditions of a point x_o to solve both problems. Our aim here is to generalize the work of Fujiwara, Han and Mangasarian in [9] to an infinite dimensional setting.

In [10] H.Maurer and J.Zowe have given a comprehensive study of this problem. We summerize their results in 3.1 to 3.10.

The set of feasible points for (P) is denoted by M, i.e. $M = G^{-1}(K)$. Let $x_o \in M$. Throughout this chapter, we assume that the first and second

Frechet derivatives $J'(x_o)$, $G'(x_o)$, $J''(x_o)$ and $G''(x_o)$ at the considered x_o exist. The sequential tangent cone $ST(M,x_o)$ and the linearizing cone $L(M,x_o)$ at x_o are used to approximate M at x_o. They are defined as :

$$ST(M,x_o) = \left\{ h \in X \mid h=\lim(x_n - x_o)/t_n,\ x_n \in M,\ t_n>0,\ t_n\to 0 \right\};$$

$$L(M,x_o) = \left\{ h \in X \mid G'(x_o)h \in K(G(x_o)) \right\}$$

$$= G'(x_o)^{-1}K(G(x_o)),$$

where $K(G(x_o)) = K + \left\{ \lambda G(x_o) \mid \lambda \in \mathbb{R} \right\}$.

Definition 3.1.

A feasible point x_o is called regular if

$$0 \in \operatorname{int}(G(x_o) + G'(x_o)X - K), \qquad (3.1)$$

where int denotes the topological interior.

Lemma 3.2.

If x_o is regular, then $L(M,x_o) \subset ST(M,x_o)$.

Theorem 3.3.

The following conditions are equivalent :

(i) $0 \in \operatorname{int}(G(x_o) + G'(x_o)X - K)$.

(ii) $0 \in \operatorname{int}(G'(x_o)X - K(G(x_o))$.

(iii) $G'(x_o)X - K(G(x_o)) = Y$.

Remark : when K is the zero cone then the definition of regularity 3.1 coinsides with definition 1.1.

Theorem 3.4. (first order necessary condition).

Let x_o be a solution for (P). If x_o is regular then

$$J'(x_o)h \geq 0 \ \forall\, h \in L(M,x_o). \qquad (3.2)$$

Theorem 3.5. (first order necessary condition).

Let x_o be a solution for (P) and let x_o be regular. Then $\exists\ l_o \in K^+$, the positive cone of K s.t.

$$J'(x_o) = G'(x_o)^* l_o \text{ and } (l_o, G(x_o)) = 0, \tag{3.3}$$

where $G'(x_o)^*$ is the adjoint induced by $G'(x_o)$ and $(\cdot, \cdot)$ is the dual pairing.

Definition 3.6.

(a) (x_o, l_o) is said to be a Kuhn-Tucker point for (P) if $x_o \in M$ and it satisfies (3.3).

(b) A functional $l_o \in K^+$ for which (3.3) holds is called the Lagrange multiplier for (P) at x_o and the function

$$F(x) = J(x) - (l_o, G(x))$$

associated with l_o is called the Lagrangian for (P) at (x_o, l_o).

Let

$$K^{l_o} = K \cap \{ y \in Y \mid (l_o, y) = 0 \}, \text{ and} \tag{3.4}$$

$$M^{l_o} = G^{-1}(K^{l_o}). \tag{3.5}$$

Since $(l_o, G(x_o)) = 0$, $G(x_o) \in K^{l_o}$ and $x_o \in M^{l_o}$. The linearizing cone $L(M^{l_o}, x_o)$ at x_o with respect to M^{l_o} is given by

$$\begin{aligned} L(M^{l_o}, x_o) &= L(M, x_o) \cap \{ h \mid (G'(x_o)^* l_o, h) = 0 \} \\ &= L(M, x_o) \cap \{ h \mid (J'(x_o), h) = 0 \}. \end{aligned} \tag{3.6}$$

This cone is the set of directions h s.t. $h \in L(M, x_o)$ for which (3.2) does not give any information. Hence we need to study the second order necessary condition about these directions.

Theorem 3.7. (second order necessary condition).

Let x_o be a solution for (P) and suppose $F(x) = J(x) - (l_o, G(x))$ is a Lagrangian for (P) at (x_o, l_o). Then

$$F''(x_o)(h,h) \geq 0 \ \forall \ h \in ST(M^{l_o}, x_o).$$

If furthermore (3.1) holds with K replaced by K^{l_o}, then

$$F''(x_o)(h,h) \geq 0 \ \forall \ h \in L(M^{l_o}, x_o). \tag{3.7}$$

In order to give a sufficient condition for (P), the linearizing cone $L(M, x_o)$ should be a "good approximation" of the feasible set M at $x_o \in M$.

Definition 3.8.

The feasible set $M = G^{-1}(K)$ is said to be approximated at $x_o \in M$ by $L(M, x_o)$, if $\exists \ h : M \to L(M, x_o)$ s.t.

$$\| h(x) - (x - x_o) \| = o(\| x - x_o \|) \ \forall x \in M.$$

H.Maurer and J.Zowe give the following conditions to guarantee that M is approximated by $L(M, x_o)$ at x_o .

Theorem 3.9.

Each of the following conditions implies that M is approximated at x_o by $L(M, x_o)$:

(i) x_o is regular, i.e. (3.1) hold.

(ii) $\dim X < \infty$ and $K(G(x_o))$ is closed.

(iii) $\dim X < \infty$, $Y = Y_1 \times \mathbb{R}^n$ and $K = \{0\} \times \mathbb{R}^n_+$, where Y_1 is a Banach space.

Theorem 3.10. (first order sufficient condition).

Let M be approximated at $x_o \in M$ by $L(M,x_o)$ and suppose that $\exists\, \beta > 0$ s.t.

$$J'(x_o)h \geq \beta \|h\| \; \forall h \in L(M,x_o).$$

Then $\exists\, \alpha > 0$ and $\rho > 0$ s.t.

$$J(x) \geq J(x_o) + \alpha \|x - x_o\| \;\; \forall x \in M \text{ with } \|x - x_o\| \leq \rho.$$

Theorem 3.11. (second order sufficient condition).

Let $x_o \in M$ and let $F(x) = J(x) - (1_o, G(x))$ be a Lagrangian for (P) at $(x_o, 1_o)$. Suppose that M is approximated at x_o by $L(M,x_o)$ and that $\exists\, \delta > 0$ and $\beta > 0$ s.t.

$$F''(x_o)(h,h) \geq \delta \|h\|^2 \; \forall h \in L(M,x_o) \cap \{h \mid (G'(x_o)^* 1_o, h) \leq \beta \|h\|\} \quad (3.8)$$

Then $\exists\, \alpha > 0$ and $\rho > 0$ s.t.

$$J(x) \geq J(x_o) + \alpha \|x - x_o\|^2 \; \forall x \in M \text{ with } \|x - x_o\| \leq \rho.$$

We consider a particular case :

(P_o) $\quad \inf J(x),$

subject to $G(x) = 0$,

where $J : X \to \mathbb{R}$,

$G : X \to Y$,

X, Y are Banach spaces.

We are going to analyse the relationship between the primal problem P_o and its Wolfe dual defined as :

(D_o) $\quad \sup J(x) - (1, G(x)),$

subject to $J'(x) - G'(x)^* 1 = 0$,

where $x \in X$, $1 \in Y^*$,

i.e. $\quad \sup \widetilde{F}(x,1) = J(x) - (1,G(x)),$

subject to $\widetilde{G}(x,1) = J'(x) - G'(x)^* 1 = 0,$

where $\tilde{F} : X \times Y^* \to R$,

$\tilde{G} : X \times Y^* \to X^*$.

We need the following definitions :

Definition 3.12.

Let X be a Banach space. Then

(a) A bilinear form B of X x X is said to be symmetric if

$$B(x,y) = B(y,x), \ \forall x,y \in X.$$

(b) A bilinear form B of X x X is said to be positive definite on a closed convex cone K if $\exists \alpha > 0$ s.t.

$$B(x,x) \geq \alpha \| x \|^2, \ \forall x \in K.$$

(c) A bilinear form B of X x X is said to be positive semidefinite on a closed convex cone K in X if

$$B(x,x) \geq 0, \ \forall x \in K.$$

Given a continuous bilinear form B on X x X, it induces an unique continuous linear mapping T_B from X into X^*. T_B is defined as

$$T_B(x)y = B(x,y) \ \forall y \in X.$$

Conversely, any linear mapping $T \in L(X,X^*)$ induces a continuous bilinear form B_T on X x X. B_T is defined as

$$B_T(x,y) = (Tx,y).$$

Definition 3.13.

A continuous bilinear form B(x,y) is said to be invertible if T_B, associated with B, is invertible. Then $T_B^{-1} \in L(X^*,X)$ and $B_{T_B^{-1}}$ is a continuous bilinear form on $X^* \times X^*$. $B_{T_B^{-1}}$ is

$$B_{T_B^{-1}}(x^*,y^*) = (y^*,T^{-1}(x^*)).$$

For simplicity, we denote $B_{T_B^{-1}}$ by B^{-1}.

Suppose x_o is a feasible regular point for problem P_o. Then the linearizing cone $L(M,x_o)$ becomes the tangent space $T(M,x_o)$, where M is the feasible set for the problem P_o, i.e.

$$\begin{aligned} L(M,x_o) &= \{ h \in X \mid G'(x_o)h \in K(G(x_o)) \} \\ &= \{ h \in x \mid G'(x_o)h = 0 \} \\ &= T(M,x_o). \end{aligned} \tag{3.9}$$

Then by remark 1.6, the annihilator $N(x_o)$ of $T(M,x_o)$ is

$$\begin{aligned} N(x_o) &= T(M,x_o)^{\perp} \\ &= \{ x^* \in X^* \mid \exists\ 1 \in Y^* \text{ s.t. } G'(x_o)^* 1 = x^* \} \\ &= G'(x_o)^* Y^*. \end{aligned} \tag{3.10}$$

Similarily, if $(x_o,1_o)$ is feasible regular point for the dual problem D_o then the linearizing cone is the kernel of $\tilde{G}'(x_o,1_o)$, i.e.

$$\begin{aligned} &\ker\tilde{G}'(x_o,1_o) \\ &= \{ (x,1) \in X \times Y^* \mid F''(x_o)x - G'(x_o)^* 1 = 0 \}, \end{aligned} \tag{3.11}$$

where $F''(x_o)$ is the second Frechet derivative of the Lagrangian at $(x_o,1_o)$ of the primal problem P_o.

Theorem 3.14.

Let x_o be a feasible regular point for primal problem P_o and let $(x_o,1_o)$ be a Kuhn-Tucker point for (P_o). Let J and G be twice continuously Frechet differentiable at x_o. Let $F''(x_o)$ be the second Frechet derivative of the Lagrangian at $(x_o,1_o)$ of the primal problem P_o. Assume that

(i) $T_{F''(x_o)} : X \to X^*$ is surjective;

(ii) $\| G'(x_o)^* 1 \| \geq a \| 1 \|\ \forall 1 \in Y^*$, $a > 0$.

Then

(i) $((x_o,1_o),0)$ is a Kuhn-Tucker point for problem D_o, where $0 \in X^{**}$;

(ii) $((x_o,1_o),0)$ satisfies the second order sufficient condition of the dual problem D_o iff $F''(x_o)$ is invertible and $F''^{-1}(x_o)$ is positive definite on the annihilator $N(x_o)$ of the tangent space $T(M,x_o)$ at x_o of the primal problem P_o.

Remark : the above second order sufficient condition refers to theorem 3.11.

Proof :

Let

$$\tilde{F}(x,l) = J(x) - (l,G(x))$$

and

$$\tilde{G}(x,l) = J'(x) - G'(x)^{*}l \ .$$

(i) Since

$$\tilde{F}_x(x_o,l_o) = J'(x_o) - G'(x_o)^{*}l_o = 0 \ ,$$

and

$$\tilde{F}_l(x_o,l_o) = -G(x_o) = 0.$$

Hence $((x_o,l_o),0)$ is a Kuhn-Tucker point for problem D_o.

(ii) Since

$$\tilde{G}_x(x_o,l_o) = F''(x_o),$$

and

$$\tilde{G}_l(x_o,l_o) = -G'(x_o)^{*}.$$

Since $F''(x_o)$ is surjective, (x_o,l_o) is regular for the dual problem D_o. Then the linearizing cone of P_o and D_o are (3.9) and (3.11) respectively. The annihilator of (3.9) is (3.10). We compute the following :

$$\tilde{F}_{xx}(x_o,l_o) = F''(x_o),$$
$$\tilde{F}_{xl}(x_o,l_o) = -G'(x_o)^{*},$$
$$\tilde{F}_{lx}(x_o,l_o) = -G'(x_o),$$

and

$$\tilde{F}_{ll}(x_o,l_o) = 0.$$

Hence, the second Frechet derivative of $\tilde{F}(x,l)$ at (x_o,l_o) is

$$\tilde{F}''(x_o,l_o) = \begin{pmatrix} F''(x_o) & -G'(x_o)^* \\ -G'(x_o) & 0 \end{pmatrix}.$$

(x_o,l_o) satisfies the second order sufficient condition for problem D_o, i.e. $\exists\, \delta < 0$ s.t.

$$\tilde{F}''(x_o,l_o)((x,l),(x,l)) \leq \delta \|(x,l)\|^2 \text{ where } (x,l) \in (3.11)$$

$$\Leftrightarrow \quad (x,l)\begin{pmatrix} F''(x_o) & -G'(x_o)^* \\ -G'(x_o) & 0 \end{pmatrix}\begin{pmatrix} x \\ l \end{pmatrix} \leq \delta \|(x,l)\|^2$$

$$\Leftrightarrow \quad (l,-G'(x_o)x) \leq \delta \|(x,l)\|^2$$

$$\Leftrightarrow \quad -(G'(x_o)^*l,x) \leq \delta \|(x,l)\|^2$$

$$\Leftrightarrow \quad (T_{F''(x_o)}x,x) \geq \delta' \|(x,l)\|^2 \quad \text{where } \delta' > 0.$$

We suppose that $F''(x_o)$ is invertible. Then

$$(T_{F''(x_o)}x,x) \geq \delta' \|(x,l)\|^2$$

$$\Leftrightarrow \quad (x^*,T^{-1}_{F''(x_o)}x^*) \geq \delta' \|(x,l)\|^2 ,$$

where $T_{F''(x_o)}\, x = x^* \in X^*$. But $T_{F''(x_o)}\, x = G'(x_o)^*l \in N(x_o)$, i.e. $\quad x^* \in N(x_o)$.

Since $F''(x_o)$ is invertible, any element $l \in Y^*$, $\exists\, x \in X$ s.t. (x,l) (3.11), i.e. x^* run through $N(x_o)$. Furthermore,

$$x^* = T_{F''(x_o)}\, x$$

$$\Rightarrow \quad \|x^*\| \leq \|T_{F''(x_o)}\|\, \|x\|,$$

and

$$x^* = G'(x_o)^*l$$

$$\Rightarrow \quad \|x^*\| \leq \|G'(x_o)^*\| \|l\|.$$

Thus $(x^*,T^{-1}_{F''(x_o)}\, x^*) \geq \beta \|x^*\|^2$ for some $\beta > 0$, $\forall\, x^* \in N(x_o)$.

Conversely, $\forall\, (x,l) \in (3.11)$,

$$x = T^{-1}_{F''(x_o)}\, x^*, \text{ where } x^* = G'(x_o)^*l \in N(x_o)$$

$$\Rightarrow \quad \|x\| \leq \|T^{-1}_{F''(x_o)}\|\, \|x^*\|.$$

By assumption, we have

$$\| 1 \| \leq a^{-1} \| G'(x_o)^* 1 \| = a^{-1} \| x^* \| .$$

Therefore,

$$(x^*, T^{-1}_{F''(x_o)} x^*) \geq \beta \| x^* \|^2 \quad , \text{ where } \beta > 0$$

$$\Rightarrow \quad (x^*, T^{-1}_{F''(x_o)} x^*) \geq \delta' \| (x,1) \|^2 , \text{ where } \delta' > 0.$$

Thus,

$$(x^*, T^{-1}_{F''(x_o)} x^*) \geq \delta' \| (x,1) \|^2 \ \forall \ (x,1) \in (3.11) \ \delta' > 0$$

$$\Leftrightarrow \quad (x^*, T^{-1}_{F''(x_o)} x^*) \geq \beta \| x^* \|^2 ; \text{ for some } \beta > 0, \ \forall \ x^* \in N(x_o).$$

It remains to show that $T_{F''(x_o)}$ is invertible. It suffices to show that $T_{F''(x_o)}$ is 1-1. Suppose that it is not,

i.e. $\tilde{x} \neq 0$ s.t. $T_{F''(x_o)} \tilde{x} = 0 \in X^*$,

$$\Rightarrow \quad (\tilde{x}, 0) \in (3.11)$$

$$\Rightarrow \quad \tilde{F}''(x_o, 1_o)((\tilde{x},0),(\tilde{x},0)) \leq \delta \| \tilde{x} \|^2 , \quad \delta < 0$$

$$\Rightarrow \quad F''(x_o)(\tilde{x},\tilde{x}) \leq \delta \| \tilde{x} \|^2 , \ \delta < 0$$

$$\Rightarrow \quad (T_{F''(x_o)} \tilde{x}, \tilde{x}) \leq \delta \| \tilde{x} \|^2 .$$

It contradicts to $(T_{F''(x_o)} \tilde{x}, \tilde{x}) = 0$. So $T_{F''(x_o)}$ is 1-1.

Theorem 3.15.

Suppose that

(i) $(x_o, 1_o)$ is a local maximun point of problem D_o ;

(ii) J, G are twice continuous Frechet differentiable at x_o .

Let $F(x) = J(x) - (1_o, G(x))$. Furthermore, assume that

(iii) $F''(x_o)$ is invertible.

Then $F''^{-1}(x_o)$ is positive semidifinite on (3.10),

i.e. $x^* = G'(x_o)^* 1 \Rightarrow (x^*, T^{-1}_{F''(x_o)} x^*) \geq 0$.

Proof :

Let $\bar{x} = G'(x_o)^*\bar{l}$.

Consider $L : X \times Y^* \times \mathbb{R} : \to X \times Y^*$. L is defined as

$$L(x,l,t) = (F_1(x,l,t), F_2(x,l,t))$$
$$= (J'(x) - G'(x)^* l \ , \ l - l_o - t\bar{l}).$$

Since $L(x_o, l_o, 0) = (J'(x_o) - G'(x_o)^* l_o \ , \ l_o - l_o) = 0$ and

$$F_{1x}(x_o, l_o, 0) = F''(x_o),$$
$$F_{1l}(x_o, l_o, 0) = -G'(x_o)^*,$$
$$F_{2x}(X_o, l_o, 0) = 0,$$

and

$$F_{2l}(x_o, l_o, 0) = I.$$

Let $\tilde{L} = \begin{pmatrix} F_{1x} & F_{1l} \\ F_{2x} & F_{2l} \end{pmatrix} = \begin{pmatrix} F''(x_o) & -G'(x_o)^* \\ 0 & I \end{pmatrix}.$

Thus , $\forall\ (\tilde{x},\tilde{l}) \in X \times Y^*, \ \exists\ (x,l) \in X \times Y^*$ s.t.

$$\tilde{L}(x,l) = \begin{pmatrix} T_{F''(x_o)}\, x - G'(x_o)^* l \\ l \end{pmatrix} = \begin{pmatrix} \tilde{x} \\ \tilde{l} \end{pmatrix},$$

where $\begin{pmatrix} x \\ l \end{pmatrix} = \begin{pmatrix} T^{-1}_{F''(x_o)}\, (\tilde{x} + G'(x_o)^*\tilde{l}) \\ \tilde{l} \end{pmatrix}$.

Clearly, $\tilde{L}$ is a linear homeomorphism of $X \times Y^*$. Hence, by implicit function theorem, $\exists\ x(t),\ l(t) \in C^1$ s.t.

(i) $x(t) \in X,\ l(t) \in Y^*,\ t \in (-\varepsilon,\varepsilon),\ \varepsilon > 0$;

(ii) $x(0) = x_o,\ l(0) = l_o$;

(iii) $L(x(t),l(t),t) = 0$,

i.e. $\quad J'(x(t)) - G'(x(t))^* l(t) = 0$ and $l(t) = l_o + t\bar{l}$.

For sufficient small t, $(x(t),\ l(t))$ is feasible for problem D_o.

Thus $J(x(t)) - (l(t),G(x(t)) \leq J(x_o) - (l_o,G(x_o))$.

Let $\theta(t) = J(x(t)) - (l(t),G(x(t))$. $\theta(t)$ achieves its local maximum at

$t = 0$. Since

$$\theta'(t) = J'(x(t))x'(t) - (l(t),G'(x(t))x'(t)) - (l'(t),G(x(t))$$
$$= -(l'(t),G(x(t)).$$

Thus,

$$\theta'(0) = -(\bar{l},G(x_o)) = 0.$$

Hence $\theta''(0) \leq 0$,

i.e. $\theta''(0)$

$$= (-l'(t),G(x(t))' \,|_{t=0}$$
$$= -(\bar{l},G'(x(t))x'(t))|_{t=0}$$
$$= -(\bar{l},G'(x_o)x'(0))$$
$$= -(G'(x_o)^*\bar{l},x'(0))$$
$$= -(\bar{x},x'(0)) \tag{3.12}$$
$$\leq 0.$$

But $J'(x(t)) - G'(x(t))^*l(t) = 0$

=> $F''(x_o)x'(0) - G'(x_o)^*\bar{l} = 0$

=> $T_{F''(x_o)}\, x'(0) = G'(x_o)^*\bar{l} = \bar{x}$

(3.12)=> $-(T_{F''(x_o)}\, x'(0),x'(0)) \leq 0$

=> $(\bar{x},T^{-1}_{F''(x_o)}\, \bar{x}) \geq 0$,

i.e. $F''^{-1}(x_o)$ is positive semidefinite on the annihilator (3.10).

Now, let X and Y be pivot Hilbert spaces.

Definition 3.16.

Let $F(x) = J(x) - \langle l_o,G(x)\rangle$, where J,G are twice continuously Frechet differentiable at x_o. Let x_o be feasible regular point for the primal problem P_o and (x_o,l_o) be a Kuhn-Tucker point for (P_o). Then

(a) (x_o,l_o) is said to satisfy the second order sufficient condition of primal problem P_o if $T_{F''(x_o)}$ is positive definite on $T(M,x_o)$, i.e. (3.9).

(b) (x_o,l_o) is said to satisfy the second order necessary condition of primal problem P_o if $T_{F''(x_o)}$ is positive semidefinite on $T(M,x_o)$, i.e. (3.9).

(c) (x_o,l_o) is said to satisfy the second order sufficient condition of the dual problem D_o if $T^{-1}_{F''(x_o)}$ is positive definite on $N(x_o)$, i.e. (3.10) and $\|G'(x_o)^* l\| \geq a\|l\| \quad \forall l \in Y^*$, $a > 0$.

(d) (x_o,l_o) is said to satisfy the second order necessary condition of the dual problem D_o if $T^{-1}_{F''(x_o)}$ is positive semidefinite on $N(x_o)$, i.e. (3.10).

Since $F''(x_o)$ is symmetric, the associated linear mapping $T_{F''(x_o)} \in L(X,X)$ is self-adjoint,

i.e. $\quad F''(x_o)(x,y) = F''(x_o)(y,x) \quad \forall x,y$ in X

$$\begin{aligned} \Rightarrow \quad \langle T_{F''(x_o)}x,y\rangle &= \langle T_{F''(x_o)}y,x\rangle \\ &= \langle T^*_{F''(x_o)}x,y\rangle \quad \forall x,y \text{ in } X \\ \Rightarrow \quad T_{F''(x_o)} &= T^*_{F''(x_o)}. \end{aligned}$$

Theorem 3.17.

Let x_o be a feasible regular point for primal problem P_o and let (x_o,l_o) be a Kuhn-Tucker point for (P_o). Let J and G be twice continuously Frechet differentiable at x_o. Let $F(x) = J(x) - \langle l_o,G(x)\rangle$. Assume that

(i) $F''(x_o)$ is invertible;

(ii) $\langle T_{F''(x_o)}x,x\rangle$ is weakly lower semicontinuous(WLSC) on $T(M,x_o)$.

If (x_o,l_o) satisfies the second order sufficient condition of primal problem P_o and satisfies the second order necessary condition of the dual problem D_o then $F''(x_o)$ is positive semidefinite on X.

Proof : the result directly follows from corollary 2.21.

Theorem 3.18.

Let x_o be a feasible regular point for primal problem P_o and let (x_o,l_o) be a Kuhn-Tucker point for (P_o). Let J and G be twice continuously Frechet differentiable at x_o. Let $F(x) = J(x) - \langle l_o,G(x)\rangle$. Assume that
(i) $F''(x_o)$ is invertible;
(ii) $\langle T^{-1}_{F''(x_o)} x,x\rangle$ is weakly lower semicontinuous on $N(x_o)$.

If (x_o,l_o) satisfies the second order sufficient condition of dual problem D_o and satisfies the second order necessary condition of primal problem P_o then $F''(x_o)$ is positive semidefinite on X.

Proof : the result directly follows from corollary 2.21.

These two theorem give us the relationship between the primal problem P_o and the dual problem D_o. Suppose that (x_o,l_o) satisfies the second order sufficient condition of primal problem. Then (x_o,l_o) solves the primal problem by theorem 3.11. We may expect that (x_o,l_o) solves the dual problem too. Then (x_o,l_o) must satisfy the second order necessary condition of the dual problem. Combining these two results, theorem 3.17 tell us that $F''(x_o)$ must be positive semidefinite on the whole space X if $F''(x_o)$ is invertible and $\langle T_{F''(x_o)} x,x\rangle$ is WLSC on $T(M,x_o)$.

Suppose that the conditions of theorem 3.18 hold. Then by theorem 3.14 (x_o,l_o) solves the dual problem D_o. We expect that (x_o,l_o) solves the primal problem P_o too. Theorem 3.18 tell us that $F''(x_o)$ must be positive semidefinite on the whole space X.

The following theorem is a sufficient condition for a point $(x_o, l_o) \in X \times Y^*$ to solve both problem P_o and D_o.

Theorem 3.19.

Let x_o be a feasible regular point for the dual problem P_o and let (x_o, l_o) be a Kuhn-Tucker point for (P_o). Let J and G be twcie continuously Frechet differentiable at x_o. Let $F(x) = J(x) - \langle l_o, G(x)\rangle$. Assume that

$$\| G'(x_o)^* l \| \geq a \| l \| \ \forall l \in Y^*, \ a > 0.$$

If $F''(x_o)$ is positive definite on whole space X then (x_o, l_o) solves the primal problem P_o and the dual problem D_o.

Proof :

Since $F''(x_o)$ is positive definite, by the corollary 2.19, $T_{F''(x_o)}$ is positive definite on $T(M, x_o)$, $T_{F''(x_o)}$ is invertible and its inverse is positive definite on $T(M, x_o)^{\perp} = N(x_o)$. Then by theorem 3.11, (x_o, l_o) solves the primal problem, and by theorem 3.14, (x_o, l_o) solves the dual problem.

Chapter 4. First and Second Order Necessary and Sufficient Conditions for Infinite Dimensional Nonlinear Programming.

Recall the primal program P in chapter 3 :

$$(P)\qquad \inf J(x),$$
$$\text{subject to } G(x) \in K,$$

where $J : X \to \mathbb{R}$, $G : X \to Y$, X and Y are Banach spaces, and K is a non-empty closed convex cone in Y. H.Maurer and J.Zowe gave a comprehensive study of this problem in [10]. They gave the first and second order necessary and sufficient conditions for this problem. We have recorded them as theorem 3.4, 3.5, 3.7 and 3.9 in chapter 3.

Now, we consider the following problem with an additional set constraint on X :

$$(P_1)\qquad \inf J(x),$$
$$\text{subject to } G(x) \in K,$$
$$x \in C,$$

where $J : X \to \mathbb{R}$, $G : X \to Y$, X and Y are Banach spaces, C is a non-empty closed convex set in X, and K is a non-empty closed convex cone in Y. We shall extend the results of H.Maurer and J.Zowe to this problem. Furthermore, we investigate the possibility of carrying out the programme developed by O.Fujiwara, S.P.Han and O.L.Mangasarian.

Throughout this chapter, we assume that the first and second Frechet derivative $J'(x_o)$, $G(x_o)$, $J''(x_o)$ and $G''(x_o)$ at the considered feasible point x_o exist.

Let

$$M_1 = C \cap G^{-1}(K). \tag{4.1}$$

$x_o \in M_1$ is called optimal for (P_1) if J restricted on M_1 assumes a local minimum at x_o. For any fixed $x \in X$ and $y \in Y$, the conical hull of $C - \{x\}$ and $K - \{y\}$ are denoted by $C_1(x)$ and $K(y)$ respectively. They are defined as :

$$C_1(x) = \left\{ \lambda(c - x) \mid c \in C,\ \lambda \geq 0 \right\} ; \tag{4.2}$$

$$K(y) = \left\{ k - \lambda y \mid k \in K,\ \lambda \geq 0 \right\} . \tag{4.3}$$

Let $x_o \in M_1$. Then the sequential tangent cone $ST(M_1,x_o)$ and the linearizing one $L(M_1,x_o)$ are defined as :

$$ST(M_1,x_o) = \left\{ h \in X \mid h = \lim(x_n - x_o)/t_n,\ x_n \in M_1,\ t_n > 0,\ t_n \to 0 \right\} ; \tag{4.4}$$

$$L(M_1,x_o) = \left\{ h \in X \mid G'(x_o)h \in K(G(x_o)),\ h \in C_1(x_o) \right\} . \tag{4.5}$$

Now, we give the more general definition of regularity.

Definition 4.1.

$x_o \in M_1$ is said to be regular if

$$0 \in \operatorname{int}\left(G'(x_o)(C - \{x_o\}) - K + G(x_o) \right) . \tag{4.6}$$

The following results 4.2 to 4.5 are due to J.Zowe and S.Kurcyusz [11].

Theorem 4.2.

Let $x_o \in M_1$. Then the following conditions are equivalent :

(i) $0 \in \operatorname{int}\left(G'(x_o)(C - \{x_o\}) - K + G(x_o) \right)$. (4.6)

(ii) $0 \in \operatorname{int}\left(G'(x_o)C_1(x_o) - K(G(x_o)) \right)$. (4.7)

(iii) $G'(x_o)C_1(x_o) - K(G(x_o)) = Y$. (4.8)

Lemma 4.3.

If x_o is regular, then $L(M_1,x_o) \subset ST(M_1,x_o)$.

Theorem 4.4. (first order necessary condition).

If the optimal solution x_o of P_1 is regular, then

$$J'(x_o)h \geq 0 \quad \forall\, h \in L(M_1, x_o) \tag{4.9}$$

Theorem 4.5. (first order necessary condition).

If the optimal solution x_o of P_1 is regular, then $\exists\, l_o \in K^+$ s.t.

(i) $J'(x_o) - G'(x_o)^* l_o \in C_1(x_o)^+$;

(ii) $(l_o, G(x_o)) = 0$.

Definition 4.6.

$l_o \in Y^*$ is called a Lagrange multiplier for (P_1) at x_o if

(i) $l_o \in K^+$, (4.10 a)

(ii) $(l_o, G(x_o)) = 0$; (4.10 b)

(iii) $J'(x_o) - G'(x_o)^* l_o \in C_1(x_o)^+$. (4.10 c)

$F(x) = J(x) - (l_o, G(x))$ is called the Lagrangian associated with l_o for (P_1) at x_o.

Suppose that l_o is a Lagrange multiplier for the optimal x_o of (P_1). Consider

$$K^{l_o} = K \cap \{ y \in Y^* \mid (l_o, y) = 0 \};$$

$$M_1^{l_o} = G^{-1}(K^{l_o}) \cap C.$$

Since $(l_o, G(x_o)) = 0$, $G(x_o) \in K^{l_o}$ and $x_o \in M_1^{l_o}$. Then the linearizing cone at x_o with respect to $M_1^{l_o}$ is

$$
\begin{aligned}
& L(M_1^{l_o}, x_o) \\
&= \{ h \in X \mid G'(x_o)h \in K^{l_o}(G(x_o)),\ h \in C_1(x_o) \} \\
&= \{ h \in X \mid G'(x_o)h = k - \lambda G(x_o),\ k \in K^{l_o},\ \lambda \geq 0,\ h \in C_1(x_o) \}
\end{aligned}
$$

$$= \left\{ h \in X \mid G'(x_o)h = k - \lambda G(x_o),\ k \in K,\ \lambda \geq 0,\ h \in C_1(x_o) \right\}$$
$$\cap \left\{ h \in X \mid (l_o, G'(x_o)h) = 0,\ h \in C_1(x_o) \right\}$$
$$= L(M_1, x_o) \cap \left\{ h \in X \mid (l_o, G'(x_o)h) = 0,\ h \in C_1(x_o) \right\}.$$

This set $L(M_1^{l_o}, x_o)$ contains the directions h which (4.9) does not give any information. So we look at the second order necessary condition.

Theorem 4.7. (second order necessary condition).

Let x_o be optimal for (P_1) and suppose $F(x) = J(x) - (l_o, G(x))$ is a Lagrangian for (P_1) at x_o. Then

$$F''(x_o)(h,h) \geq 0 \ \forall h \in ST(M_1^{l_o}, x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o).$$

If furthermore (4.6) holds with K replaced by K^{l_o}, then

$$F''(x_o)(h,h) \geq 0 \ \forall h \in L(M_1^{l_o}, x_o) \cap \left\{ h \in X \mid J'(x_o)h = 0,\ h \in C_1(x_o) \right\}.$$

Proof :

$\forall\ x \in M_1^{l_o}$,

$\Rightarrow x \in C$ and $G(x) \in K^{l_o}$

$\Rightarrow x \in C$, $G(x) \in K$ and $(l_o, G(x)) = 0$

$\Rightarrow x$ is a feasible for (P_1).

Furthermore, $F(x) = J(x) - (l_o, G(x)) = J(x)$ if $x \in M_1^{l_o}$. Thus, $F(x)$ assumes a local minimun at x_o. Therefore, $F'(x_o)h \geq 0,\ \forall h \in ST(M_1^{l_o}, x_o)$, i.e.

$$F'(x_o)h = (J'(x_o) - G'(x_o)^* l_o, h) \geq 0,\ \forall h \in ST(M_1^{l_o}, x_o).$$

So,

$$F'(x_o)h = 0,\ \forall h \in ST(M_1^{l_o}, x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o).$$

Hence,

$$F''(x_o)(h,h) \geq 0,\ \forall h \in ST(M_1^{l_o}, x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o).$$

If (4.6) holds with K replaced by K^{l_o}, then by lemma 4.3, $L(M_1^{l_o}, x_o)$ is a subset of $ST(M_1^{l_o}, x_o)$. So

$$L(M_1^{l_o},x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o)$$
$$\subseteq ST(M_1^{l_o},x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o).$$

By theorem 4.5, $J'(x_o) - G'(x_o)^* l_o \in C_1(x_o)^+$. If

$$\tilde{h} \in L(M_1^{l_o},x_o) \cap \{ h \in X \mid J'(x_o)h = 0,\ h \in C_1(x_o) \},$$

then

$$(J'(x_o) - G'(x_o)^* l_o, \tilde{h}) \geq 0,$$

i.e. $\quad (G'(x_o)^* l_o, \tilde{h}) \leq 0.$

So, $\quad (G'(x_o)^* l_o, \tilde{h}) = 0.$

Thus, $\tilde{h} \in \ker(J'(x_o) - G'(x_o)^* l_o)$.

Therefore $\quad L(M_1^{l_o},x_o) \cap \{ h \in X \mid J'(x_o)h = 0,\ h \in C_1(x_o) \}$

$$\subseteq L(M_1^{l_o},x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o)$$
$$\subseteq ST(M_1^{l_o},x_o) \cap \ker(J'(x_o) - G'(x_o)^* l_o).$$

Hence $\quad \forall\ h \in L(M_1^{l_o},x_o) \cap \{ h \in X \mid J'(x_o)h = 0,\ h \in C_1(x_o) \}$,

$$F''(x_o)(h,h) \geq 0.$$

Now, we are in the position to develop some sufficient conditions for the problem P_1.

Definition 4.8.

The feasible set M_1 is said to be approximated at $x_o \in M_1$ by $L(M_1,x_o)$, if there exists a map $h : M_1 \to L(M_1,x_o)$ s.t.

$$\| h(x) - (x - x_o) \| = o(\| x - x_o \|)\ \forall\ x \in M_1.$$

Theorem 4.9. (first order sufficient condition).

Let M_1 be approximated at $x_o \in M_1$ by $L(M_1,x_o)$ and suppose that $\exists\ \beta > 0$ s.t.

$$J'(x_o)h \geq \beta \|h\|\ \ \forall h \in L(M_1,x_o).$$

Then $\exists \alpha > 0$, $\rho > 0$ s.t. $J(x) \geq J(x_o) + \alpha \| x - x_o \|$ $\forall x \in M_1$ with $\| x - x_o \| \leq \rho$.

Proof :

For simplicity we assume that $x_o = 0$.

Since M_1 is approximated at 0 by $L(M\ ,0)$, $\forall x \in M_1$, x can be decomposited as $x = h(x) + z(x)$ where $h(x) \in L(M_1,0)$ and $\|z(x)\| = o(\|x\|)$.

Consider

$$\begin{aligned} & J(x) - J(0) \\ &= J'(0)x + r(x) \quad \text{where } |r(x)| = o(\|x\|) \\ &= J'(0)h(x) + J'(0)z(x) + r(x) \\ &\geq \beta\|h(x)\| + r_1(x) \end{aligned} \tag{4.11}$$

where $r_1(x) = J'(0)z(x) + r(x)$. Therefore, $|r_1(x)| = o(\|x\|)$. We can choose $\rho > 0$ so small that

$$\| z(x)\| < \tfrac{1}{2}\|x\|, \ |r_1(x)| \leq \tfrac{1}{4}\beta\|x\|$$

for $x \in M_1$, $\|x\| < \rho$. Hence

$$\|h(x)\| = \|x - z(x)\| \geq \tfrac{1}{2}\|x\|,$$

and from (4.11), we have

$$J(x) - J(0) \geq \tfrac{1}{2}\beta\|x\| - \tfrac{1}{4}\beta\|x\| \geq \tfrac{1}{4}\beta\|x\|.$$

We have our desired results.

Lemma 4.10.

Let $h : M_1 \to L(M_1,x_o)$ be a map s.t.

$$\|h(x) - (x - x_o)\| = o(\|x - x_o\|) \ \forall x \in M_1.$$

Then $\forall \gamma > 0$, $\exists \rho > 0$ s.t. $\|h(x) - (x - x_o)\| \leq \gamma\|h(x)\|$ $\forall x \in M_1$ with $\| x - x_o\| \leq \rho$.

Proof :

Choose $\delta \in (0,1)$, $\exists\ \rho > 0$ s.t. $\forall x \in M_1$ with $\| x - x_o \| \leq \rho$ one has $\| h(x) - (x - x_o) \| \leq \delta \| x - x_o \|$ and thus

$$\begin{aligned} \| h(x) \| &= \| h(x) - (x - x_o) + (x - x_o) \| \\ &\geq | \| h(x) - (x - x_o) \| - \| x - x_o \| | \\ &= \| x - x_o \| - \| h(x) - (x - x_o) \| \\ &\geq (1-\delta) \| x - x_o \|. \end{aligned}$$

Hence $\forall x \in M_1$ with $\| x - x_o \| \leq \rho$ and $x \neq x_o$, then $h(x) \neq 0$.

$$\frac{\| h(x) - (x - x_o) \|}{\| h(x) \|} = \frac{\| h(x) - (x - x_o) \|}{\| x - x_o \|} \times \frac{\| x - x_o \|}{\| h(x) \|} \leq \frac{\delta}{1 - \delta}.$$

for all $x \in M_1 \setminus \{x_o\}$ and $\| x - x_o \| \leq \rho$. For any given $\gamma > 0$, we can choose $\delta \in (0,1)$ s.t.

$$\frac{\delta}{1 - \delta} < \gamma$$

Hence the result follows.

Lemma 4.11. (lemma 5.5 of [10]).

Let B be a continuous symmetric bilinear form on X x X, H a subset of X and $\delta > 0$ with

$$B(h,h) \geq \delta \|h\|^2 \text{ for all } h \in H.$$

Then there are $\delta_o > 0$ and $\gamma > 0$ s.t.

$$B(h + z, h + z) \geq \delta_o \| h + z \|^2 \text{ for all } h \in H,\ z \in X \text{ and } \|z\| \leq \gamma \|h\| .$$

Theorem 4.12. (second order sufficient condition).

Let $x_o \in M_1$ and let $F(x) = J(x) - (l_o, G(x))$ be a Lagrangian for (P_1) at x_o. Suppose that M_1 is approximated at x_o by $L(M_1, x_o)$ and that there are $\delta > 0$ and $\beta > 0$ s.t.

$$F''(x_o)(h,h) \geq \delta \| h \|^2 \text{ for all } h \in L(M_1,x_o) \cap \{ h | (1_o,G'(x_o)h) \leq \beta \| h \| \}.$$

Then there exist $\alpha > 0$ and $\rho > 0$ s.t. $J(x) \geq J(x_o) + \alpha \| x - x_o \|^2$ for all $x \in M_1$ with $\| x - x_o \| \leq \rho$.

Proof :

For simplicity we assume $x_o = 0$.

Since M_1 is approximated at 0 by $L(M_1,0)$, for all $x \in M_1$ can be written as

$$x = h(x) + z(x),\ h(x) \in L(M_1,0),\ \| z(x) \| = o(\| x \|).$$

We claim that there exist $\alpha_1 > 0$, $\alpha_2 > 0$, $\gamma_1 > 0$ and $\gamma_2 > 0$ s.t.

$$J(x) \geq J(0) + \alpha_1 \| x \|^2 \text{ whenever } x = h(x) + z(x) \in M_1,\ \| x \| \leq \rho_1$$
$$\text{and } (1_o,G'(0)h(x)) \leq \beta \| h(x) \| ,$$

and

$$J(x) \geq J(0) + \alpha_2 \| x \| \text{ whenever } x = h(x) + z(x) \in M_1,\ \| x \| \leq \rho_2$$
$$\text{and } (1_o,G'(0)h(x)) > \beta \| h(x) \|.$$

By the definition of first Frechet-derivative one has for all $x \in M_1$:

$$F(x) = F(0) + F'(0)x + \tfrac{1}{2}F''(0)(x,x) + r(x),$$

i.e.

$$J(x) - (1_o,G(x))$$
$$= J(0) - (1_o,G(0)) + (J'(0) - G'(0)^{*}1_o,x) + \tfrac{1}{2}F''(0)(x,x) + r(x).$$

Since $\forall\ x \in M_1$

$$\Rightarrow x \in C$$
$$\Rightarrow x \in C_1(0)$$
$$\Rightarrow (J'(0) - G'(0)^{*}1_o,x) \geq 0 \text{ for } J'(0) - G'(0)^{*}1_o \in C_1(0)^{+},$$

thus

$$J(x) \geq J(x) - (1_o,G(x))$$
$$\geq J(0) + \tfrac{1}{2}F''(0)(x,x) + r(x).$$

Let $H = L(M_1,0) \cap \{ h \mid (1_o,G'(0)h) \leq \beta \| h \| \}$. Since $F''(0)(h,h) \geq \delta \| h \|^2$ for all $h \in H$, by lemma 4.11, there exist $\delta_0 > 0$ and $\gamma > 0$ s.t.

$$F''(0)(h + z, h + z) \geq \delta_0 \| h + z\|^2 \text{ for all } h \in H,\ z \in X \text{ and } \| z\| \leq \gamma \| h\|.$$

By lemma 4.10, for such γ, there exists $\rho_1 > 0$ s.t.

$$\| z(x)\| \leq \gamma \| h(x)\| \quad \text{whenever } \|x\| \leq \rho_1 ,\ x \in M_1, \tag{4.12}$$

i.e. $\exists\ \delta_0 > 0,\ \rho_1 > 0$ s.t.

$$\tfrac{1}{2}F''(0)(x,x) \geq \delta_0 \| x \|^2 \text{ whenever } x \in M_1,\ \|x\| \leq \rho_1 \text{ and } x = h(x) + z(x),\ h(x) \in H.$$

Moreover, we can choose ρ_1 small enough such that

$$|r(x)| \leq \tfrac{1}{4}\delta_0 \| x \|^2 \qquad \text{for} \| x \| \leq \rho_1 .$$

Then $\qquad J(x) \geq J(0) + \frac{1}{2}\delta_0 \| x \|^2 - \frac{1}{4}\delta_0 \| x \|^2,$

i.e. $\qquad J(x) \geq J(0) + \frac{1}{4}\delta_0 \| x \|^2$ whenever $x = h(x) + z(x) \in M_1$, and $h(x) \in H$, $\|x\| \leq \rho_1$.

Now, we consider the case $x = h(x) + z(x) \in M_1, \| x \| \leq \rho_1$ but $(1_o, G'(0)h(x)) > \beta \| h(x) \|$. By the definition of first Frechet-derivative one has for all $x \in M_1$:

$$J(x) - J(0) = J'(0)h(x) + J'(0)z(x) + r(x).$$

Since $J'(0) - G'(0)^* 1_o \in C_1(0)^+$, $(J'(0) - G'(0)^* 1_o ,h(x)) \geq 0$ for all $h(x)$ in $L(M_1,0)$. So $J'(0)h(x) \geq (1_o, G'(0)h(x)) \geq \beta \| h(x) \|$. Hence

$$J(x) - J(0) \geq \beta \| h(x) \| + r_1(x), \tag{4.13}$$

where $r_1(x) = J'(0)z(x) + r(x)$. Since $r_1(x) = J'(0)z(x) + r(x)$ and $\|z(x) \| = o(\|x\|)$, $|r_1 (x)| = o(\|x\|)$. By (4.12),

$$\| x\| \leq \| h(x) \| + \| z(x) \| \leq (1 + \gamma)\| h(x)\| .$$

(4.13) becomes

$$J(x) - J(0) \geq \frac{\beta}{1+\gamma} \| x\| + r_1(x).$$

We can choose ρ_2 sufficient small, $\rho_2 \in (0, \rho_1]$ s.t. there exists $\alpha_2 > 0$

$$J(x) - J(0) \geq \alpha_2 \| x \| \text{ whenever } x = h(x) + z(x) \in M_1 \text{ and}$$
$$\| x \| \leq \rho_2 \leq \rho_1 \text{ , } (1 \text{ ,} G'(0)h(x)) > \beta \| h(x) \|.$$

We have proved our claim.

We can choose $\rho_1 < 1$ then $\|x\|^2 \leq \|x\|$ whenever $\|x\| \leq \rho_2 \leq \rho_1 < 1$. Hence $0 < \alpha = \min \{\alpha_1, \alpha_2\}$, $0 < \rho = \rho_2 < 1$,

$$J(x) \geq J(0) + \alpha \|x\|^2 \text{ for all } x \in M_1 \text{ with } \|x\| \leq \rho .$$

The following theorem gives conditions for which M_1 is approximated at x_o by $L(M_1, x_o)$.

Theorem 4.13.

Each of the following conditions implies that M_1 is approximated at x_o by $L(M_1, x_o)$:

(i) x_o is regular.

(ii) $\dim X < \infty$ and $K(G(x_o))$ is closed.

(iii) $\dim X < \infty$, $Y = Y_1 \times \mathbb{R}^n$ and $K = \{0\} \times \mathbb{R}^n_+$, where Y_1 is a Banach space.

Proof :

(i) For $x \in M_1$, $x \neq x_o$, we have

$$G(x) - G(x_o) = G'(x_o)(x - x_o) + r(x, x_o), \tag{4.14}$$

where $\|r(x,x_o)\| = o(\|x - x_o\|)$. Since x_o is regular, by theorem 4.2

$$G'(x_o)C_1(x_o) - K(G(x_o)) = Y.$$

Then we can apply theorem 2.1 in [11], there exists $\alpha > 0$ s.t.

$$B_Y \subseteq \alpha(G'(x_o)(C - \{x_o\})_1 - (K - \{G(x_o)\})_1),$$

where B_Y is the unit ball in Y,

$$(C - \{x_o\})_1 = (C - \{x_o\}) \cap B_X, \ B_X \text{ is the unit ball in } X;$$
$$(K - \{G(x_o)\})_1 = (K - \{G(x_o)\}) \cap B_Y.$$

Then there exists $z(x) \in \alpha \|r(x,x_o)\| B_X \cap C_1(x_o)$, and $k(x) \in K(G(x_o))$ s.t.

$$r(x,x_o) = G'(x_o)z(x) - k(x). \tag{4.15}$$

We set $h(x) = (x - x_o) + z(x) \in C_1(x_o)$, since $C_1(x_o)$ is convex. Then

$$\begin{aligned} &\|h(x) - (x - x_o)\| \\ &= o(\|z(x)\|) \\ &= o(\|r(x,x_o)\|) \\ &= o(\|x - x_o\|), \end{aligned}$$

and

$$\begin{aligned} &G'(x_o)h(x) \\ &= G'(x_o)((x - x_o) + z(x)) \\ &= G'(x_o)(x - x_o) + G'(x_o)z(x) \\ &= G'(x_o)(x - x_o) + r(x,x_o) + k(x) && \text{by (4.15)} \\ &= G(x) - G(x_o) + k(x) && \text{by (4.14)} \\ &= G(x) - G(x_o) + k_1 - \lambda G(x_o) && \text{where } k_1 \in K \text{ since } k(x) \in K(G(x_o)) \\ &= G(x) + k_1 - (1 + \lambda)G(x_o) \\ &\in K(G(x_o)) \quad \text{since } K \text{ is convex, } G(x) + k_1 \in K. \end{aligned}$$

Hence M_1 is approximated at x_o by $L(M_1,x_o)$.

(ii) Now suppose $\dim X < \infty$ and $K(G(x_o))$ is closed. For simplicity we assume that $x_o = 0$. For every $x \in M_1$, select $h(x) \in L(M_1,x_o)$ s.t.

$$\|h(x) - x\| = \min\left\{\|h - x\| \mid h \in L(M_1,0),\ \|h\| \leq 2\|x\|\right\}.$$

For $L(M_1,0) = G'^{-1}(0)(K(G(0)) \cap C_1(0)$, $L(M_1,0)$ is closed since $K(G(0))$ and $C_1(0)$ are closed. Therefore, $\left\{h \mid h \in L(M_1,0),\ \|h\| \leq 2\|x\|\right\}$ is compact. Thus $h(x)$ exists.

It remains to prove that $h(x) = o(\|x\|)$. Suppose that it is not true. Then there exists $\varepsilon > 0$ and a sequence of points x_n in M_1, $x_n \neq 0$, s.t. for $h_n = h(x_n)$:

$$\|x_n\| \leq 1/n \text{ and } \|h_n - x_n\| > \varepsilon \|x_n\| \tag{4.16}$$

As $\dim X < \infty$, we may assume that the sequence $\frac{x_n}{\|x_n\|}$ converges to some $\bar{h} \in X$ with $\|\bar{h}\| = 1$. Then by the definition of $ST(M_1,0)$, $\bar{h} \in ST(M_1,0)$. We claim that $ST(M_1,0) \subset L(M_1,0)$ if $K(G(0))$ is closed. Let $d \in ST(M_1,0)$, i.e. $\exists \{y_n\} \subset M_1$ and $\lambda_n > 0$, $\lambda_n \in \mathbb{R}$ s.t. $y_n/\lambda_n \to d$. Since $C_1(0)$ is closed, $d \in C_1(0)$. Moreover, $G(y_n) \in K$. Then $(G(y_n) - G(0))/\lambda_n \in K(G(0))$.

$$\frac{G(y_n) - G(0)}{\lambda_n} \to G'(0)d$$

As $K(G(0))$ is closed, $G'(0)d \in K(G(0))$. By definition of $L(M_1,0)$, $d \in L(M_1,0)$. Hence $\bar{h} \in L(M_1,0)$. Let

$$z_n = \bar{h} - x_n/\|x_n\|.$$

We get $x_n + \|x_n\| z_n = \|x_n\|\bar{h} \in L(M_1,0)$ and $\|x_n + \|x_n\| z_n\| = \|x_n\| \leq 1/n$.

$$\begin{aligned} &\| h_n - x_n \| \\ &= \min\left\{ \|y - x_n\| \mid y \in L(M_1,0), \|y\| \leq 2\|x_n\| \right\} \\ &\leq \|x_n + \|x_n\| z_n - x_n\| \\ &= \|x_n\| \, \|z_n\| \to 0 \text{ as } z_n \to 0. \end{aligned}$$

It contradicts to (4.16). This proves (ii).

(iii) Since $K(G(x_o))$ is closed for $K = \{0\} \times \mathbb{R}^n_+$, by (ii), M_1 is approximated at x_o by $L(M_1,x_o)$.

For problem (P), we defined its Wolfe dual problem D as follow :

(D) $\qquad \sup J(x) - (1,G(x)),$

$\qquad$ subject to $J'(x) - G'(x)^* 1 = 0,$

where $J : X \to \mathbb{R}$, $G : X \to Y$, X and Y are Banach spaces, and K^+ is the positive cone of the closed convex cone K in Y, i.e.

$$\sup F_D(x,1) = J(x) - (1,G(x)),$$

$$\text{subject to } H(x,1) = J'(x) - G'(x)^* 1 = 0, \ 1 \in K^+.$$

The feasible set is $M_D = (X \times K^+) \cap H^{-1}(0)$. Suppose (x_o, l_o) is a Kuhn-Tucker point for the primal problem (P). The conical hull of $X \times K^+ - \{(x_o, l_o)\}$ is

$$C_D(x_o, l_o) = \{ r((x,l) - (x_o, l_o)) \mid (x,l) \in X \times K^+, r \geq 0 \}$$
$$= X \times K^+(l_o),$$

where $K^+(l_o) = K^+ - \{ r l_o \mid r \geq 0 \}$. The linearizing cone of M_D at (x_o, l_o) is

$$L(M_D, (x_o, l_o)) = \{ (x,l) \in C_D(x_o, l_o) \mid F''(x_o)x - G'(x_o)^* l = 0 \},$$

where $F''(x_o)$ is the second Frechet-derivative of the Lagrangian of the primal problem. In order to carry out the programme developed by O.Fujiwara, S.-P.Han and O.L.Mangasarian as in theorem 3.14, it requires that $G'(x_o)^*(K^+(l_o))$ coinsides with the positive cone of the linearizing cone used in the primal problem (P). However, the linearizing cone of the primal problem now is

$$L(M, x_o) = \{ h \in X \mid G'(x_o)h \in K(G(x_o)) \},$$

where $K(G(x_o)) = K + \{ r G(x_o) \mid r \in R \}$.

Thus in general, such requirement cannot be met. Since $-\frac{1}{2} l_o \in K^+(l_o)$, for all h in $L(M, x_o)$,

$$-\tfrac{1}{2}(G'(x_o)^* l_o, h)$$
$$= -\tfrac{1}{2}(l_o, G'(x_o)h)$$
$$= -\tfrac{1}{2}(l_o, k - rG(x_o)) \quad \text{where } r \geq 0,\ k \in K$$
$$= -\tfrac{1}{2}(l_o, k) \quad \text{since } (l_o, G(x_o)) = 0$$
$$\leq 0.$$

Reference

[1] A.D.Ioffe and V.M.Tihomirov, Theory of Extremal Problems(Studies in Mathematics and Its Applications, 6, North-Holland, Amsterdam, New York, Oxford ,1979).

[2] Jean-Pierre Aubin, Applied Functional Analysis(John Wiley, New York, 1979).

[3] Lorch, Spectral Theory(New York . Oxford University Press . 1962).

[4] Bachman Narici, Functional Analysis(Academic Press, New York and London, 1972).

[5] S.-P.Han and O.L.Mangasarian, "A conjugate decomposition of the Euclidean space", Proc. Natl. Acad. Sci. USA Vol 80(1983) pp.5156 - 5157, Mathematics.

[6] S.-P.Han and O.L.Mangasarian, "Conjugate cone characterization of positive definite and semidefinite matrices", Linear Algebra and its application 56(1984) pp.89 - 103.

[7] M.Seetharama Gowda, "Cone characterizations of positive semidefinite operators on a Hilbert space", Linear Algebra and its applications 64(1985) pp.77 - 83.

[8] Andrzej P.Wierzbicki and Stanislaw Kurcyusz, "Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in Hilbert space", SIAM J.Control and Optimization Vol 15, No. 1, 1977, pp.25 -56.

[9] O.Fujiwara, S.-P.Han and O.L.Mangasarian, "Local duality of nonlinear programs", SIAM J.Control and Optimization, Vol 22, No.1, 1984, pp.162-169.

[10] H.Maurer and J.Zowe, "First and second-order necessary and sufficient conditions for infinite-dimensional programming problems", Mathematical programming 16, 1979, pp.98 - 110.

[11] J.Zowe and S.Kurcyusz, "Regularity and stability for the mathematical programming problem in Banach spaces", Appl. Math. Optim. 5, 1979, pp.49 - 62.

[12] Helmut Maurer, "First and second order sufficient optimality conditions in mathematical programming and optimal control", Mathematical Programming Study 14, 1981, pp.163 - 177.

[13] A. Wouk, A Course of Applied Functional Analysis(John Wiley, New York, 1979).

[14] T.Kobayaski, "Initial state determination for distributed parameter systems", SIAM, Journal on Control and Optimization, 14(1976), pp.934-944.